EXTENDED EXPERIMENTAL PROCEDURES

Cell Culture

Human breast cancer cell lines (BT549, CAMA1, HCC1187, HCC1569, HCC38, MDA-MB-231, MDA-MB-415, MDA-MB-468, T47D, ZR-75-1) were cultured in RPMI 1640 medium supplemented with 10% fetal bovine serum (FBS) and 1% penicillin/streptomycin. For cells stably introduced with tetracyclin-inducible genes/shRNAs, Tet-approved FBS (Clontech) was used. For packaging virus, HEK293T cells were grown in DMEM with 10% FBS and 1% penicillin/streptomycin.

Primary Breast Tumor Culture

Tumors were minced, resuspended in Collagenase/Hyaluronidase (Stemcell Technologies, Cat# 07912), and incubated at 37°C with rotation for 2 to 4 hours. After centrifugation at 100 x g for 2 min, cells were seeded in the DMEM/F-12 medium supplemented with 5% FBS, 1% penicillin/streptomycin, 20 ng/ml EGF, 0.5 mg/ml hydrocortisone, 100 ng/ml cholera toxin, and 10 μg/ml insulin. After 1-2 d of culture, cells were trypsinized, washed, and seeded in 96-well plates for drug treatment.

Plasmids

To generate pLKO-tet-on-shRNAs targeting CDK7, oligonucleotides were designed and synthesized (IDT). Following annealination, double-stranded oligonucleotides were directly ligated with tet-on-pLKO vector (21915, Addgene; Wiederschain et al., 2009) that was digested with AgeI and EcoRI. The sequences are

Scramble (fwd: **ccgg**gtggactcttgaaagtactat**ctcgag**atagtactttcaagagtccac**ttttt**g

rv: **aattaaaaa**gtggactcttgaaagtactat**ctcgag**atagtactttcaagagtccac),

CDK7_1 (fwd: **ccgg**gctgacaccatcacacatcaa**ctcgag**ttgatgtgtgatggtgtcagc**ttttt**g

rv: **aattcaaaaa**gctgacaccatcacacatcaa**ctcgag**ttgatgtgtgatggtgtcagc),

CDK7_2 (fwd: **ccgg**gctgtagaagtgagtttgtaa**ctcgag**ttacaaactcacttctacagc**ttttttg;**

rv: aattcaaaaagctgtagaagtgagtttgtaa**ctcgag**ttacaaactcacttctacagc),

CDK7_3 (fwd: **ccgg**gcaggagacgacttactagat**ctcgag**atctagtaagtcgtctcctgc**ttttttg**

rv: **aattcaaaaa**gcaggagacgacttactagat**ctcgag**atctagtaagtcgtctcctgc).

Construction of lenti-CRISPR/CAS9 vectors targeting transcriptional CDK was performed following the protocol associated with the backbone vector (49535, Addgene; Shalem et al., 2014;). The following sequences were chosen from the published libraries (Shalem et al., 2014; Wang et al., 2014), with priority given to sequences that match the early coding exons of targeted genes. The non-bold sequence is gene-specific.

GFP (fwd: **CACCG**GGGCGAGGAGCTGTTCACCG;

rv: **AAAC**CGGTGAACAGCTCCTCGCCC**C**),

CDK7_1 (fwd: **CACCG**GAAGCTGGACTTCCTTGGGG;

rv: **AAAC**CCCCAAGGAAGTCCAGCTTC**C**),

CDK7_2 (fwd: **CACCG**ATCTCTGGCCTTGTAAACGG;

rv: **AAAC**CCGTTTACAAGGCCAGAGAT**C**),

CDK8 (fwd: **CACCG**CGAGGACCTGTTTGAATACG

rv: AAACCGTATTCAAACAGGTCCTCG***C***),

CDK9 (fwd: ***CACCG***GCACCGCAAGACCGGCCAGA

rv:AAACTCTGGCCGGTCTTGCGGTGC***C***),

CDK12 (fwd: **CACCG**GGGGGAGACAGATCTCCACC

rv: **AAA**CGGTGGAGATCTGTCTCCCCC**C**),

CDK13 (fwd: **CACCG**AGGAGCGGCAACAGCAGCGG;

rv:**AAAC**CCGCTGCTGTTGCCGCTCCT**C**),

CDK19 (fwd: **CACCG**ATTATGCAGAGCATGACTTG;

rv: **AAAC**CAAGTCATGCTCTGCATAAT**C**).

EGFR_1 (fwd: **CACCG**GCAACGTGGAGAGCATCCAG;

rv: **AAAC**CTGGATGCTCTCCACGTTGC**C**).

EGFR_2 (fwd: **CACCG**ATCATAATTCCTCTGCACAT;

rv: **AAAC** ATGTGCAGAGGAATTATGAT**C**).

EGFR_3 (fwd: **CACCG**AGGCACGAGTAACAAGCTCA;

rv: **AAAC** AGGCACGAGTAACAAGCTCA**C**).

EN1 (fwd: **CACCG**ACTCGCTCTCGTCTTTGTCC;

rv: **AAAC**GGACAAAGACGAGAGCGAGT**C**).

ETS1 (fwd: **CACCG**GACATCATTTCTTTGCTGCT;

rv: **AAAC**AGCAGCAAAGAAATGATGTC**C**).

FosL1 (fwd: **CACCG**CTTCCTCCGGTTCCTGCACT;

rv: **AAAC**AGTGCAGGAACCGGAGGAAG**C**).

FoxC1 (fwd: **CACCG**GGAGTGGTGCCCTACCTCGG;

rv: **AAAC**CCGAGGTAGGGCACCACTCC**C**).

MYC (fwd: **CACCG**GCCGTATTTCTACTGCGACG;

rv: **AAAC**CGTCGCAGTAGAAATACGGC**C**).

SOX9_1 (fwd: **CACCG** TTCACCGACTTCCTCCGCCG;

rv: **AAAC**CGGCGGAGGAAGTCGGTGAA**C**).

SOX9_2 (fwd: **CACCG**CGTGTTCTCGGTGTCCGAGC;

rv: **AAAC**GCTCGGACACCGAGAACACG**C**).

TWIST1 (fwd: **CACCG**CTGTCGTCGGCCGGCGAGAC;

rv: **AAAC**GTCTCGCCGGCCGACGACAG**C**).

Virus Infection

Lentiviruses were generated in HEK293T cells by transfecting cells with packaging DNA plus tet-on-pLKO or lenti-CRISPR vectors. Typically 2 μg vector DNA, 1.5 μg pCMV-dR8.91, and 0.5 μg pMD2-VSVG, 12 μl lipid of Metafectene Pro (Biontex) were used; DNA and lipid were pre-diluted in 300 μl PBS individually and then mixed. After 15 min of incubation, the DNA-lipid mixtures were added to HEK293T cells ($3x10^6$ cells seeded in one T-25 flask one day earlier). Viral supernatant was collected two and three days after transfection, filtered through 0.45-μm membranes, and added to target cells in the presence of polybrene (8 μg/ml, Millipore). Puromycin (1.5 μg/ml) was used to treat cells for two days for selection, which eliminated all cells in an uninfected control group.

Cell Proliferation Assay

For 96-well plate assay, cells were plated at the density of 2000 cells per well, and on the next day treated with THZ1 or THZ2 of various concentrations. After 48-hour incubation, cells were fixed and stained with crystal violet. The staining was then extracted by adding each well with 10% acetic acid, with absorbance measured at 590 nm with 750 nm as a reference.

For cells with tet-on-shRNA, cells plated in 12-well plates ($1–2 \times 10^4$) in total 1 ml medium. On the next day, each well was added with 110 μl medium without or with 1 μg/ml doxycycline (to reach a final concentration of 100 ng/ml). The addition of doxycycline was repeated every two days. One week after the initial doxycycline treatment, cells were fixed and viewed by microscope.

For cells introduced with lenti-CRISPR vectors, after puromycin selection, cells were harvested and seeded in 12-well plates. Cells were fixed 7-10 days after seeding, and stained with crystal violet. The plates were washed extensively, and imaged with a

flatbed scanner. For quantification of the staining, 1 ml 10% acetic acid was added to each well to extract the dye. The absorbance was measured at 590 nm with 750 nm as a reference.

Animal Studies

All animal experiments were conducted in accordance with the animal use guidelines from the National Institutes of Health and with protocols approved by the Dana-Farber Cancer Institute Animal Care and Use Committee. Nude mice (CrTac:NCr-Foxn1nu) were γ-irradiated with a single dose of 400 rads six hours before transplantation of cells. Breast cancer cells were harvested and resupended in 40% Matrigel-Basement Membrane Matrix, LDEV-free (BD Biosciences), and then injected (100 µl per site) into the fourth pair of mammary fat pads of mice. Tumors were measured in two dimensions by using manual calipers. Tumor volume was calculated using the formula: $V = 0.5 \times \text{length} \times \text{width} \times \text{width}$. Animal with tumor established (mean tumor volume of ~200 mm^3) were randomly divided into two groups, which were then treated with vehicle (10% DMSO in D5W, 5% dextrose in water) or THZ2 (3 mg/ml, prepared in vehicle solutions) at the dose of 10 mg/kg intraperitoneally twice daily. Tumor volume was measure every 2-3 days. Upon harvesting tumors, tumors were cut into half, with one half fixed in formalin overnight and then in 70% ethanol for histopathology analysis, and the other half snap frozen in liquid nitrogen for immunoblotting.

Live Cell Imaging

Time-lapse imaging of cancer cells expressing H2B-GFP was performed on a Nikon Ti motorized inverted microscope, equipped with a perfect focus system and a humidified incubation chamber (37°C, 5% CO2) (Nikon Imaging Center, Harvard Medical School). Cells were seeded in 24-well glass-bottom plate, and on the next day treated with

vehicle control (0.1% DMSO, v/v) or THZ1 (50 nM). Imaging started right after the treatment and lasted for 24 hours, with images captured every 5 min using a 20x object lens. Images were analyzed using ImageJ (National Institute of Health).

Immunoblotting

Cultured cells were lysed in RIPA buffer (50 mM Tris, pH 7.4, 150 mM NaCl, 1% Nonidet P-40, 0.5% sodium deoxycholate, and 0.1% sodium dodecyl sulfate), which was supplemented with protease inhibitors cocktail (Roche) and phosphatase inhibitors cocktail (Thermo Scientific). After incubation, lysates were cleared by centrifugation and then subjected to protein concentration assay (BCA kit, Thermo Scientific). Tumor samples were homogenized in RIPA buffer supplemented with protease/phosphatease inhibitors using Bullet blender (Next advance). Lysates were then processed as above. Twenty microgram of lysates were used for SDS-PAGE. Nitrocellulose membrane with protein transferred was blocked with 5% non-fat milk and was then incubated with primary antibodies overnight at 4°C. After washing, the membrane was incubated with fluorophore-conjugated secondary antibodies for one hour at room temperature. The membrane was then washed and scanned with an Odyssey Infrared scanner (Li-Cor Biosciences).

Primary antibodies used were anti-phospho-CTD-RNAPII-S2 (04-1571, Millipore), anti-phospho-CTD-RNAPII-S5 (04-1572, Millipore), anti-phospho-CTD-RNAPII-S7 (04-1570, Millipore), anti-RNAPII (A300-653A, Bethyl), anti-CDK7 (2916, Cell Signaling Technology), anti-CDK8 (51-8130-HR, BD Pharmingen), anti-CDK9(A303-493A, Bethyl), anti-CDK13 (A301-458A, Bethyl), anti-CDK19 (HPA008053, Sigma), anti-cleaved PARP (Asp214) (9541, Cell Signaling Technology), anti-cleaved Caspase 3 (Asp175) (9664, Cell Signaling Technology), anti-Cyclin B1 (4135, Cell Signaling Technology), anti-

Aurora B (3094, Cell Signaling Technology), anti-MELK (ab108529, Abcam), and anti-α-tubulin (Sigma). Secondary antibodies used were Alexa680-conjugated anti-rabbit IgG (Invitrogen) and IRDye800-conjugated anti-mouse IgG (Rockland).

Cell Cycle Analysis

Cells were harvested by trypsinization, and fixed by 70% ethanol (−20°C) (added dropwise while vortexing the samples). Cells were then stained with solutions containing propidium iodide (50 μg/ml, Sigma) and DNase-free RNase A (50 μg/ml, Sigma). After 30 min of incubation, the samples were washed and re-suspended in PBS. The samples were run on a LSRFortessa (BD Biosciences), with single cells gated via plotting FL3-A to FL3-H to exclude cell debris and doublets.

Ion Torrent Library preparation, Sequencing, and Data Analysis

cDNA libraries were constructed using the Ion AmpliSeq Transcriptome Human Gene Expression Kit (A26325, Life Technologies), and were sequenced using the Ion Proton System according to manufacturer instructions. Briefly, 10 ng of total RNA samples were used for the preparation of each cDNA library. Eight libraries were multiplexed and clonally amplified to obtain template-positive ion sphere particles by using the Ion OneTouch 2 System (Life Technologies), and were sequenced on an Ion Torrent Proton using one PI chip kit V2. The sequencing run generated 9.9 G of raw data with more than 91 million reads. Data were analyzed by Torrent Suite 4.4 and ampliSeqRNA analysis plugin (Life Technologies), to identify differentially expressed genes.

RNA Extraction and Synthetic RNA Spike-In.

Cells were treated for 6 h with vehicle control or THZ1, and then subjected to total RNA extraction using RNeasy Mini kit (Qiagen), with QIAshredder spin column for

homogenization and an on-column DNase digestion (79254, Qiagen). Total RNA was spiked-in with ERCC RNA Spike-In Mix (Ambion, 4456740) (Lovén et al., 2012), treated with DNA-freeTM DNase I (Ambion, AM1906) and analyzed on Agilent 2100 Bioanalyzer for integrity. RNA with the RNA Integrity Number (RIN) above 9.8 was hybridized to GeneChip PrimeView Human Gene Expression Arrays (Affymetrix).

Microarray Sample Preparation and Analysis

For microarray analysis, 100 ng of total RNA containing ERCC RNA Spike-In Mix (see above) was used to prepare biotinylated aRNA (cRNA) according to the manufacturer's protocol (3' IVT Express Kit, Affymetrix 901228). Briefly, total RNA undergoes T7 oligo(dT)-primed reverse transcription to synthesize first-strand cDNA containing a T7 promoter sequence. This cDNA is then converted into a double-stranded DNA template for transcription using DNA Polymerase and RNase H to simultaneously degrade the RNA and synthesize second strand cDNA. In vitro transcription synthesizes aRNA and incorporates a biotin-conjugated nucleotide. The aRNA is then purified to remove unincorporated NTPs, salts, enzymes, and inorganic phosphate. Fragmentation of the biotin-labeled aRNA prepares the sample for hybridization onto GeneChip 3' expression arrays. Samples were prepared for hybridization using 12.5 µg of biotinylated aRNA in a 1X hybridization cocktail according the Affymetrix hybridization manual. Additional hybridization cocktail components were provided in the Affymetrix GeneChip Hybridization, Wash and Stain Kit. GeneChip arrays (Human PrimeView, Affymetrix 901837) were hybridized in a GeneChip Hybridization Oven at 45°C for 16 hrs at 60 RPM. Washing was done using a GeneChip Fluidics Station 450 according to the manufacturer's instructions, using the buffers provided in the Affymetrix GeneChip Hybridization, Wash and Stain Kit. Images were extracted with Affymetrix GeneChip Command Console (AGCC), and analyzed using Expression Console. A Primeview CDF

that included probe information for the ERCC controls (GPL16043), provided by Affymetrix, was used to generate .CEL files. We processed the CEL files as described in Kwiatkwoski et al., 2014 using standard tools available within the affy package in R. The CEL files were processed with the expresso command to convert the raw probe intensities to probeset expression values. The parameters of the expresso command were set to generate Affymetrix MAS5-normalized probeset values. We used a *loess* regression to re-normalize these MAS5 normalized probeset values, using only the spike-in probesets to fit the *loess*. The affy package provides a function, *loess* normalize, which will perform *loess* regression on a matrix of values (defined using the parameter mat) and allows for the user to specify which subset of data to use when fitting the *loess* (defined using the parameter subset, see the affy package documentation for further details). For this application the parameters mat and subset were set as the MAS5-normalized values and the row-indices of the ERCC control probesets, respectively. The default settings for all other parameters were used. The result of this was a matrix of probeset expression values normalized to the control ERCC probes. Normalized probeset intensity values are converted into log2 space after adding a pseudocount of 1. The log2 probeset values from duplicates were averaged together. RefSeq transcript-level expression values are determined by selecting the probeset for each transcript with the highest mean intensity across all experiments. Log2 fold-changes per transcript are calculated by subtracting log2 DMSO (untreated) from log2 drug-treated values. These post-processing values for each sample are included in Supplemental Table 2. Transcripts are considered expressed if their DMSO expression value is > log2(100). Heatmaps in Figure 5A and Supplementary Figure 5A were made on all expressed transcripts using heatmap.2.

Chromatin Immunoprecipitation

Cells were cross-linked for 10 min at room temperature by the addition of one-tenth of the volume of 11% formaldehyde solution (11% formaldehyde, 50mM HEPES pH 7.4, 100mM NaCl, 1mM EDTA pH 8.0, 0.5mM EGTA pH 8.0) to the growth media followed by 5 min quenching with 1/20 the volume of growth medium of 2.5M glycine. Cells were washed twice with PBS, then the supernatant was aspirated and the cell pellet was flash frozen in liquid nitrogen. Frozen crosslinked cells were stored at -80°C. 50 µL of Dynal magnetic beads (Sigma) were blocked with 0.5% BSA (w/v) in PBS. Magnetic beads were bound with 10 µg of the indicated antibody. For H3K27Ac occupied genomic regions, we performed ChIP-Seq experiments using an Abcam (AB4729A) antibody. The affinity-purified antibody was raised in rabbit against an epitope corresponding to amino acids 1-100 of human Histone H3 that is acetylated at K27. Crosslinked cells were lysed with lysis buffer 1 (50mM HEPES pH 7.5, 140mM NaCl, 1mM EDTA, 10% glycerol, 0.5% NP-40, and 0.25% Triton X-100), pelleted and resuspended in lysis buffer 2 (10 mM TrisHCl pH 8.0, 200 mM NaCl, 1 mM EDTA, 0.5 mM EGTA). The subsequent pellet was resuspended in and sonicated in sonication buffer (50 mM HEPES pH 7.5, 140 mM NaCl, 1 mM EDTA pH 8.0, 1 mM EGTA, 0.1% Na-deoxycholate, 0.1% SDS, and 1% Triton X-100). Cells were sonicated for 10 cycles at 30 s each on ice (18-21 W) with 60 s on ice between cycles. Sonicated lysates were cleared and incubated overnight at 4°C with magnetic beads bound with antibody to enrich for DNA fragments bound by the indicated factor. Beads were washed two times with sonication buffer, one time with sonication buffer with 500 mM NaCl, one time with LiCl wash buffer (10 mM TrisHCl pH 8.0, 1 mM EDTA, 250 mM LiCl, 0.5% NP-40, 0.5% Na-deoxycholate) and one time with TE. DNA was eluted in elution buffer (50 mM TrisHCl pH 8.0, 10 mM EDTA, 1% SDS). Cross-links were reversed overnight. RNA and protein were digested using RNase A and Proteinase K, respectively and DNA was purified with phenol chloroform extraction and ethanol precipitation.

ChIP-Seq Analysis

Illumina sequencing libraries generated and data was processed according to Lin et al., 2012. In brief, libraries were generated for ChIP samples following the Illumina TruSeq™ DNA Sample Preparation v2 kit protocol with minor changes. Briefly, these minor changes included 1) adapters from the genomic DNA prep kit were diluted 1:50 prior to use and 2) the samples were size selected after PCR amplification, rather than before. All ChIP-Seq data sets were aligned using bowtie 1.0.1 (Langmead et al., 2009) to build NCBI36/hg19 of the human genome with parameters –p 4, -k 2, -m 2,–best, –sam. Wiggle files for gene tracks were created using Macs 1.4.2 (Zhang et al., 2008) with options –w –S –space=50 –nomodel –keep-dup=all to count reads in 50bp bins, and were displayed in the UCSC genome browser.

Super-Enhancer Identification and Assignment

Super-enhancers were largely identified as described by Chipumoro et al., 2014 using the ROSE2 software described in Loven et al., 2013 and uploaded to github.com/BradnerLab/pipeline. macs 1.4.2 20120305 was used to identify regions statistically enriched in H3K27ac ChIP-Seq reads with –p 1e-9 and input DNA control. Peaks were called twice using parameters –keep-dup=1 and –keep-dup=auto to partially account for focal amplifications in tumor genomes. The union of these peaks was used as constituent enhancers for ROSE2. ROSE2 was used to identify super-enhancers and separate them from typical enhancers in a mathematical manner and was run with parameters –s 12500, -t 2000, -g hg19 to rank enhancers by H3K27ac signal minus input DNA control signal. Stitched enhancers were assigned to the active transcript most proximal to the center of the enhancer; active transcripts for this analysis were identified as having promoters in the top 2/3 of H3K27ac signal within a region +/-500bp

from the TSS. Density calculations were performed using bamToGFF (github.com/BradnerLab/pipeline) with parameters –e 200, -m 1, -r, -d.

Motif analysis

Enrichments of occurrences of DNA sequence motifs in regions (Figure 5F) were determined using AME (Buske et al., 2010, PMID: 20147307). The sequences in constituent enhancers of super-enhancers associated with genes in the Achilles cluster were compared to the whole pool of super-enhancer constituents with parameters –p-value-threshold 0.0002 –scoring totalhits –method fisher. Two motif position-weight matrix files were created containing MA0523.1 (TCF7L2) and MA0144.2 (STAT3) from Jaspar core motifs 2014 and SMAD3_DBD (SMAD3) from Jolma 2013; and CTCF_full (CTCF) from Jolma 2013.

Gene Set Enrichment Analysis.

Gene Set Enrichment Analysis (Mootha et al., 2003, PMID: 12808457) was performed using a gene list pre-ranked by fold-change upon THZ1 treatment from patient -derived TNBC primary cells (Figure 5G, Supplementary Table 3). The 166 Achilles cluster genes - established from TNBC cell lines - was used as the gene set to test for enrichment of the Achilles cluster genes among THZ1 –sensitive genes in TNBC primary cells.

REFERENCES

Chipumuro, E., Marco, E., Christensen, C.L., Kwiatkowski, N., Zhang, T., Hatheway, C.M., Abraham, B.J., Sharma, B., Yeung, C., Altabef, A., et al. (2014). CDK7 inhibition suppresses super-enhancer-linked oncogenic transcription in MYCN-driven cancer. Cell 159, 1126-1139.

Christensen, C.L., Kwiatkowski, N., Abraham, B.J., Carretero, J., Al-Shahrour, F., Zhang, T., Chipumuro, E., Herter-Sprie, G.S., Akbay, E.A., Altabef, A., et al. (2014). Targeting Transcriptional Addictions in Small Cell Lung Cancer with a Covalent CDK7 Inhibitor. Cancer cell 26, 909-922.

Kwiatkowski, N., Zhang, T., Rahl, P.B., Abraham, B.J., Reddy, J., Ficarro, S.B., Dastur, A., Amzallag, A., Ramaswamy, S., Tesar, B., et al. (2014). Targeting transcription regulation in cancer with a covalent CDK7 inhibitor. Nature 511, 616-620.

Langmead, B., Trapnell, C., Pop, M., and Salzberg, S.L. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome biology 10, R25.

Lin, C.Y., Loven, J., Rahl, P.B., Paranal, R.M., Burge, C.B., Bradner, J.E., Lee, T.I., and Young, R.A. (2012). Transcriptional amplification in tumor cells with elevated c-Myc. Cell 151, 56-67.

Loven, J., Hoke, H.A., Lin, C.Y., Lau, A., Orlando, D.A., Vakoc, C.R., Bradner, J.E., Lee, T.I., and Young, R.A. (2013). Selective inhibition of tumor oncogenes by disruption of super-enhancers. Cell 153, 320-334.

Shalem, O., Sanjana, N.E., Hartenian, E., Shi, X., Scott, D.A., Mikkelsen, T.S., Heckl, D., Ebert, B.L., Root, D.E., Doench, J.G., et al. (2014). Genome-scale CRISPR-Cas9 knockout screening in human cells. Science 343, 84-87.

Wang, T., Wei, J.J., Sabatini, D.M., and Lander, E.S. (2014). Genetic screens in human cells using the CRISPR-Cas9 system. Science 343, 80-84.

Wiederschain, D., Wee, S., Chen, L., Loo, A., Yang, G., Huang, A., Chen, Y., Caponigro, G., Yao, Y.M., Lengauer, C., et al. (2009). Single-vector inducible lentiviral RNAi system for oncology target validation. Cell cycle 8, 498-504.

Zhang, Y., Liu, T., Meyer, C.A., Eeckhoute, J., Johnson, D.S., Bernstein, B.E., Nusbaum, C., Myers, R.M., Brown, M., Li, W., et al. (2008). Model-based analysis of ChIP-Seq (MACS). Genome biology 9, R137.